DNA Profiling Analysis of Dust and Source Material Collected in the CRPAQS Fall 2000 Sampling Campaign

DRAFT FINAL REPORT
CONTRACT NO. 2000-08PM

Prepared for:
California Air Resources Board
1102 Q Street
P.O. Box 2815
Sacramento, CA 95812

Submitted by:
Department of Land, Air, and Water Resources
One Shields Avenue
University of California, Davis
Davis, CA 95616

Prepared by:

Kate M. Scow	**Principal Investigator**
Mara J. Johnson	**Post-doctoral researcher**
Geoff N. Elliott	**Post-doctoral researcher**

The statements and conclusions in this Report are those of the contractor and not necessarily those of the California Air Resources Board. The mention of commercial products, their source, or their use in connection with material reported herein is not to be construed as actual or implied endorsement of such products.

ACKNOWLEDGMENTS

This Report was submitted in fulfillment of Air Resources Board Contract No. 2000-08PM "DNA profiling analysis of dust and source material collected in the CRPAQS fall 2000 sampling campaign" by the University of California at Davis under the sponsorship of the California Air Resources Board. Work was completed as of September 2002.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1.0 PROJECT SUMMARY

1.1 Background and Statement of the Problem

Levels of PM_{10} (particulate matter $\leq$10 μm aerodynamic diameter) in California's Central Valley regularly exceed both State and Federal air quality standards. Particularly in late summer and early fall, soil-derived dust constitutes the dominant fraction of PM_{10} and may be generated by agricultural operations, vehicular traffic on paved and unpaved roads, and construction activities. These principle sources of dust are indistinguishable by conventional analytical methods, prompting basic research for alternate methods of source characterization, including the use of biological markers derived from soil microorganisms (bacteria, fungi, and protozoa). We have developed fingerprinting methods based on the direct extraction and analysis of nucleic acids (DNA) associated with soil microbial communities. DNA analysis focuses on the genetic code contained within cell nuclei. The principle advantages of DNA analysis include low detection limits and the potential use of highly specific gene sequences. This study presents the results of the first field application of these methods with the objective of characterizing PM_{10} dust and potential sources of dust during cotton harvest in Corcoran, California.

1.2 Project Objectives

i) Collect and characterize (using DNA fingerprinting) fugitive PM_{10} dust from a range of collection sites within the CRPAQS network in Corcoran, CA during the fall sampling period.
ii) Collect and characterize (using DNA fingerprinting) potential source material within and adjacent to the CRPAQS fall network.
iii) Determine relationships among potential source materials and between dust and potential source materials.

1.3. Project Conclusions

1. We were successful in extracting and analyzing DNA from PM_{10} collected on Hi-Vol sampler filters at the Corcoran stationary site and from filters collected in the field during agricultural operations. The filters at the stationary site, which collected PM_{10} over 2 day sampling periods, contained masses of dust from 120 to 390 mg. These amounts are sufficient to obtain enough DNA to permit PCR-based fingerprint analysis.
2. The DNA fingerprints of dust on filters changed over time and varied with different peak PM_{10} events. Thus, there was no unique dust fingerprint common to all dust samples. Dust samples were more similar to their potential sources than to one another. This provided support for the idea that DNA fingerprinting is potentially useful for relating PM_{10} to potential sources.
3. In some, but not all cases, there were similarities between dust samples collected at the Corcoran stationary site and from some of the sources. For some of the samples, a subset, but not all, bands within their fingerprints appeared to be the same. In some cases, source samples upwind of the Corcoran stationary site were not collected because the wind direction did not follow predicted patterns. In all cases, there was no opportunity for replication in sample collection and thus hindered the ability to detect small differences and to draw statistically-based conclusions.

4. We collected a library of DNA fingerprints for potential sources samples from agricultural soils and unpaved roads surrounding Corcoran. We could differentiate source samples from one another, but there did not appear to be a strong relationship between source fingerprint and land use or crop type (in contrast to findings in our previous ARB contract).

2.0 SAMPLE COLLECTION

2.1 Calendar of Field and Laboratory Efforts

July 21, 2000	ARB meeting in Sacramento to discuss Fall Intensive Monitoring Campaign.
August 16, 2000	Planned for meeting at Corcoran to evaluate Corcoran core site for installation of additional Hi-Vol sampler. (Contact: Chuck McDade)
August 23, 2000	Gathered information and farm manager contacts for agricultural source sampling in vicinity of Corcoran. (Contacts: Jim Sweet and Chuck McDade)
September 5, 2000	Secured location for additional Hi-Vol sampler at Corcoran core site. (Contact: Chuck McDade)
September 8, 2000	Received notification that T&B Systems would change Hi-Vol filters.
September 18, 2000	Initiated agricultural source sampling. (Contact: Roger Isom)
September 27, 2000	Assembled Hi-Vol sampler at Corcoran core site.
October 27-November 7, 2000	Conducted mobile filter and source sample collection at Corcoran.
November 2000-March 2001	Extracted DNA.
March 2001-May 2001	Performed PCR and gel electrophoresis
June 2001-March 2003	Conducted data analysis and wrote report

2.2 Sample Collection

2.2.1 Source soil samples

The locations of source samples were identified after discussions with ARB staff. After securing permission from Dennis Tristao (J. G. Boswell), Rista Gilkey (Gilkey Farms), Trent Hair (Hall Ranch), Michael Boyett (Boyett Farms), and Gary O'Neil (Hansen Farms) samples were collected from 0-10 cm soil depth from agricultural fields and unpaved roads from areas within 1-3 miles of Corcoran city limits. Based on previous information about prevailing wind direction, most source sampling occurred in areas northwest and west of Corcoran. Samples also were collected south of Corcoran, but not due north or east of Corcoran. Two types of samples were collected. Composite source soil samples were collected approximately 50 meters apart from four to five locations within a designated section identified for each field. For each location, five subsamples were collected from a central position at the site, and at four positions in orthogonal directions 10 meters from the center position (north, east, south, and west). These samples were mixed to create a composite sample. Composite unpaved road samples were collected from the surface with a dustpan and broom. For each unpaved road

sample, five subsamples were collected at 10 meter intervals and mixed to create a composite sample.

2.2.2 Ambient air samples-agricultural operations

The primary goal of this objective was to obtain an adequate mass of material to test whether microbial fingerprints specific to particular soils could be detected in field-collected PM_{10} derived from those soils during agricultural operations. Sterilized (by autoclave) 6" x 8" quartz Q-MA filters were pre-weighed in the laboratory, inserted into sterile, polypropylene bags, and sealed until use in the field. Unused blank filters were used as controls. Ambient air PM_{10} samples were collected during two agricultural operations with a Hi-Vol sampler (Model SA/G 1200 Anderson Instruments Inc. Smyrna, GA) mounted on the bed of a pick-up truck. The Hi-Vol sampler was positioned downwind of the field operation and switched on when the dust plume reached the Hi-Vol sampler. In addition, the truck was repositioned to maintain alignment with the dust plume. Samples were collected for approximately 60 minutes.

2.2.3 Ambient air samples-Corcoran Core Site

The primary goal of this objective was to analyze dust samples collected within Corcoran, at sites representative of ARB monitoring locations, under standard sampling conditions. Ambient air PM_{10} samples were collected with a Hi-Vol sampler (Model SA/G 1200 Anderson Instruments Inc. Smyrna, GA) on a building adjacent to the Corcoran COP Core Site. Sterilized (by autoclave) 6" x 8" quartz Q-MA filters were pre-weighed in the laboratory, inserted into sterile, polypropylene bags, and sealed until use in the field. Unused blank filters were used as negative controls. T&B Systems retrieved filters with ambient air PM_{10} samples at 8 am after a 2-3 day period of sampling. The sampling interval was selected to collect sufficient dust to provide a large enough sample for analysis and was terminated during rainfall. Information recorded for each sample included flow rate, temperature, wind speed and direction, and ambient air temperature, both at the start and end of each collection interval. The mass of PM_{10} collected on each filter was determined by weighing samples in the laboratory.

3.0 SAMPLE ANALYSIS

3.1 DNA Extraction and Analysis

A DNA fingerprinting method was used to characterize soil and dust samples. As dust is derived from soil, and soils harbor specialized microbial communities as a function of specific environmental influences (such as moisture and nutrient availability), DNA fingerprinting methods can characterize a sample based on its microbial community composition. Soil communities include bacteria (prokaryotes), fungi (eucaryotes), protozoa (eucaryotes), and other organisms, all of which contain biochemical material that can be extracted and analyzed. The types and amounts of extracted biochemical material not only comprise a fingerprint of a sample, but also constitute a set of multivariate data applicable in classification and multivariate analyses. A previous project with the Air Resources Board (Contract # 94-321) set the foundation for using biological tools to characterize dust and sources of dust.

We utilized a polymerase chain reaction (PCR) DNA-based method, which has a very low detection limit, because we anticipated small sample sizes of field-collected dust. Our project used PCR primers designed to amplify a small portion of the last section of the small subunit ribosomal RNA (rRNA) gene, the first portion of the large-subunit rRNA gene, and the internal transcribed spacer (ITS) region between these two genes. Then the amplified sequences are separated by length to yield a DNA fingerprint of the microbial community within each sample(Garcia-Martinez et al., 1999; Jensen et al., 1993). The output of a DNA fingerprint is an electrophoresis gel containing multiple bands that represent major groups of microorganisms living in a particular soil sample. The location and number of these bands can be compared across samples to determine percent similarity among source samples and between dust and source samples.

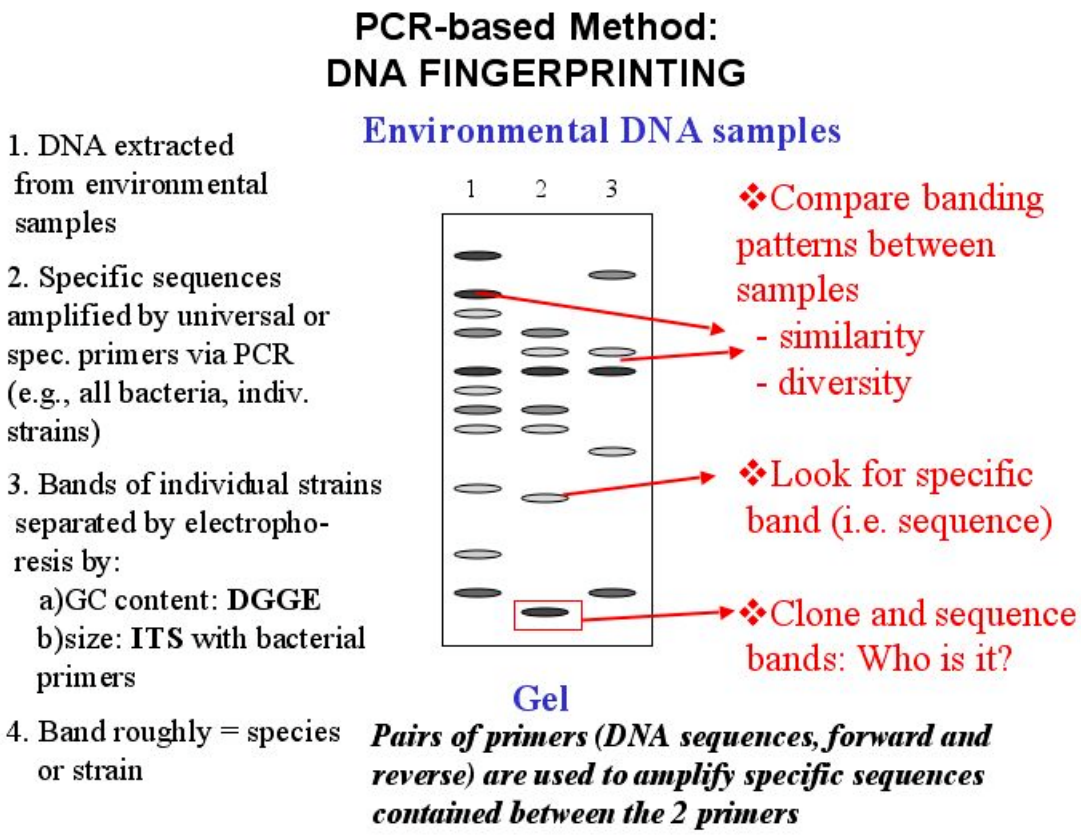
DNA from microbial communities was obtained by subjecting soil and dust samples to chemical and physical treatments which lysed (break open) microbial cells and released microbial DNA into solution. Following DNA purification, the DNA was quantified (only for soil extracts, which have higher DNA content) with a spectrophotometer to standardize concentrations in preparation for the polymerase chain reaction (PCR). The protocol is presented in Appendix A.

The products of the PCR, multiple copies of ITS fragments (bands) from whole-community DNA, range in size from 300 to 1400 base pairs. These fragments were separated in a polyacrylamide gel matrix on the basis of fragment length and visualized with DNA-binding stain. The varied position and intensity of the bands constitute the DNA fingerprint. The gels were photographed to capture a digital image of the DNA fingerprint pattern. The pattern, which looks similar to a bar code, is the genetic profile that constitutes the data used for distinguishing sources and for soil-dust comparisons. All profiles were analyzed using cluster analysis to allow inferences to be made about sample similarities and relationships. Digital images of DNA fingerprints were imported to GelCompar II, a software package designed for DNA fingerprint analysis (Applied Maths, Kortrijk, Belgium). Images were processed as recommended by the GelCompar II designers with reference to image processing points from Rademaker and De Bruijn (1997). Details are provided in Appendix B.

3.2 Statistical Methods

Figure 1 is a conceptual diagram of a typical DNA fingerprint on an electrophoresis gel and the type of information that can be obtained from such a fingerprint. To analyze the fingerprint data, individual lanes (columns) in the polyacrylamide gels containing bacterial PCR products (horizontal bands) were scanned and stored, by sample, using the GelCompar II software. Each lane was digitized into a 320 pixel densitometric curve that comprised a fingerprint pattern. Fingerprints were compared with a whole pattern analysis method using the Pearson product moment correlation method (Pearson, 1926), which directly compares samples based on densitometric curve data. The Pearson product moment correlation was previously demonstrated as a powerful method for analysis of complex DNA fingerprints (Haene et al., 1993; Rademaker and De Bruijn, 1997). Cluster diagrams were generated by the unweighted pair group method, using arithmetic averages (UPGMA) algorithm (Sneath and Sokal, 1973), to determine sample groupings.

Figure 1 Conceptual Diagram of DNA Fingerprinting



4.0 RESULTS AND DISCUSSION

4.1 Sample Collection

4.1.1 Source soils and labeling

Eighty agricultural source and 12 unpaved roadway samples were collected (Table 1) from a total of 19 sections surrounding Corcoran (Figure 2). The source samples are divided into three groups by the first letter in the sample name. Samples beginning with the letter "C" were collected within three miles or less of Corcoran. Samples beginning with the letter "SW" were collected in a seven mile transect southwest of Corcoran. Samples beginning with the letter "X" were collected intensively at particular locations. The number in the sample name designates the section from which the sample was taken. The last letters (e.g. A, B, C) refer to positions in the sections from which the samples were taken. This nomenclature also designates locations of unpaved road (upr) samples. In DNA fingerprint figures, the additional numbers following the sample names represent laboratory replicates. Cotton, wheat, alfalfa, hay, and oat crops in various stages of growth, harvest, and post-harvest tillage were represented in the sample set. Unfortunately, heavy rainfall during the period our research team was in Corcoran precluded the collection of comparable source samples within city limits.

4.1.2 Agricultural field dust

Two paired PM_{10} and source samples were collected with the mobile collection unit during agricultural tillage operations. The first was collected during the second disking of cotton stubble in the southeast corner of section 21, southwest of Corcoran. The source sample, C21F, was a composite sample that represented the area that was disked during operation of the Hi-Vol sampler. The quartz filter (Q601) held 120 mg of PM_{10}, which was an adequate mass for DNA fingerprinting. The second was collected during the final disking of a cotton field in the northeast corner of section 12, west of Corcoran. The source sample, X12, was a composite as described above. The quartz filter (Q602) held 110 mg of PM_{10}, also a mass adequate for DNA fingerprinting. Filter blanks were negative for DNA. After a period of heavy rainfall, the emissions from agricultural operations were greatly reduced and further agricultural field dust sampling could not be conducted.

Table 1. Source sample[a] inventory and collection details.

Sample	Crop	Field state	Date collected	Likely soil classification of sample
C4A	Cotton	Unpicked	10/27/2000	175
C4B	Cotton	Unpicked	10/27/2000	113
C4D	Cotton	Unpicked	10/27/2000	101
C4F	Cotton	Picked, not shredded	10/27/2000	134/101
C4G	Cotton	Picked, not shredded	10/27/2000	101
C8	Wheat	Furrowed (green)	10/27/2000	101/168
C9A	Wheat	Winter prepared	10/27/2000	134
C9C	Wheat	Winter prepared	10/27/2000	101
C9D	Wheat	Winter prepared	10/27/2000	101
C9E	Wheat	Winter prepared	10/27/2000	101
C10A	Cotton	Unpicked	10/27/2000	134/119/175
C10B	Cotton	Unpicked	10/27/2000	134
C16A(w)	Cotton	Winter prepared	10/27/2000	101/119
C16B(w)	Cotton	Winter prepared	10/27/2000	101/119/167/168
C17A	Wheat	Winter prepared	10/27/2000	101
C17B	Wheat	Winter prepared	10/27/2000	101/168
C17C	Wheat	Winter prepared	10/27/2000	101
C17D	Cotton	Disked once	10/27/2000	101/168
C20B	Cotton	Unpicked	10/28/2000	101/116
C20C	Wheat	Furrowed (green)	10/28/2000	168
C20E	Cotton	Unpicked	10/28/2000	116
C20F	Wheat	Furrowed (green)	10/28/2000	116/168
C21B	Wheat	Furrowed (green)	10/28/2000	101/168
C21C	Cotton	Flattened	10/28/2000	134/101
C21E	Wheat	Furrowed (green)	10/28/2000	101/168
C21F	Cotton	Disked	10/28/2000	101
C26B	Cotton	Unpicked	10/28/2000	101/119/134/168/124
C26C	Cotton	Unpicked	10/28/2000	101/153/134
C26E	Cotton	Unpicked	10/28/2000	101/124
C26F	Cotton	Unpicked	10/28/2000	119/124/101
C27C	Cotton	Unpicked	10/28/2000	168/134
C27D	Cotton	Unpicked	10/28/2000	168
C27E	Cotton	Unpicked	10/28/2000	168/124
C27F	Cotton	Unpicked	10/28/2000	168/124
C28B	Cotton	Shredded	10/28/2000	116/124/101
C28E	Hay	Disked and manured	10/28/2000	116
C28F	Cotton	Unpicked	10/28/2000	116/124
C31/32B	Cotton	Unpicked	10/27/2000	153/101
C31/32C	Cotton	Unpicked	10/27/2000	153/101
C31/32D	Cotton	Unpicked	10/27/2000	101
C31/32E	Cotton	Unpicked	10/27/2000	101
C31/32F	Cotton	Unpicked	10/27/2000	101/168
C32/33B	Oats?	Cut not disked	10/27/2000	unknown
C32/33C	Oats?	Cut not disked	10/27/2000	unknown
C32/33D	Oats?	Cut not disked	10/27/2000	unknown
C32/33E	Oats?	Cut not disked	10/27/2000	unknown
C32/33F	Oats?	Cut not disked	10/27/2000	unknown
C33B	Wheat	Disked	10/27/2000	117, 134, 113, 175
C33E	Wheat	Disked	10/27/2000	117, 134, 113, 175

[a] The first letter designations for the samples include C = Corcoran, SW = Southwest transect, X = Intensive Sampling.
The number designations refer to the Section from which the sample was collected.
The final letter designates the position in the field from which the sample was collected.

Table 1 continued. Source sample[a] inventory and collection details.

Sample	Crop	Field state	Date collected	Likely soil classification of sample
SW3A	Cotton	Picked, not shredded	10/28/2000	163
SW3B	Cotton	Unpicked	10/28/2000	163
SW3C	Cotton	Unpicked	10/28/2000	163
SW3D	Cotton	Unpicked	10/28/2000	163
SW9A	Cotton	Unpicked	10/28/2000	163
SW9B	Cotton	Picked, not shredded	10/28/2000	163
SW9C	Cotton	Picked, not shredded	10/28/2000	163
SW19A	Cotton	Picked, not shredded	10/28/2000	101
SW19B	Cotton	Picked, not shredded	10/28/2000	116
SW19C	Alfalfa	Growing	10/28/2000	116
SW19D	Alfalfa	Growing	10/28/2000	116
SW25A	Cotton	Flattened	10/28/2000	163/116
SW25B	Cotton	Flattened	10/28/2000	163
SW25C	Alfalfa	Harvested and disked	10/28/2000	163
SW25D	Alfalfa	Harvested and disked	10/28/2000	163
SW35A	Cotton	Picked, not shredded	10/28/2000	163
SW35B	Cotton	Picked, not shredded	10/28/2000	163
SW35C	Cotton	Picked, not shredded	10/28/2000	163
SW35D	Cotton	Picked, not shredded	10/28/2000	163
C9UPR	Road	(adjacent winter prepared wheat)	10/27/2000	upr (101, 119, 175)
C16UPR	Road	(adjacent winter prepared cotton)	10/30/2000	upr (101, 119, 168, 167)
C20UPR	Road	adjacent (furrowed wheat(green))	10/30/2000	upr (116, 168, 101)
C26UPR	Road	(adjacent cotton - unpicked)	10/30/2000	upr (101, 119, 124, 168, 134)
C28UPR	Road	(adjacent cotton - various)	10/28/2000	upr (116, 124, 101)
C31/32UPR	Road	(adjacent cotton - unpicked)	10/27/2000	upr (101, 153, 168)
C32/33UPR	Road	(adjacent oats - cut not disked)	10/27/2000	upr (101, 168, 117, 134, 153)
SW19UPR	Road	(adjacent alfalfa/cotton)	10/28/2000	upr (116, 154, 101, 168)
X12UPR	Road	(adjacent cotton - disked)	10/30/2000	upr (101, 153, 168)
X18AUPR	Road	(adjacent alfalfa/cotton)	10/30/2000	upr (101, 116, 124)
X18BUPR	Road	(adjacent alfalfa/cotton)	10/30/2000	upr (101, 116, 124)
X34UPR	Road	(cotton)	10/30/2000	upr (116, 168, 101, 124)
X12A1	Cotton	Disked	10/25/2000	153
X12A3	Cotton	Disked	10/25/2000	153
X12A5	Cotton	Disked	10/25/2000	153
X12A7	Cotton	Disked	10/25/2000	153
X12X	Cotton	Disked	10/24/2000	153
X12	Cotton	Disked	10/25/2000	101/153
X11T7N	Hay	Disked	10/25/2000	168/101
X11T7S	Hay	Disked	10/25/2000	168/101
X11T8N	Hay	Disked	10/25/2000	168/101
X11T8S	Hay	Disked	10/25/2000	168/101
X11T9N	Hay	Disked	10/25/2000	168/101
X11T9S	Hay	Disked	10/25/2000	168/101

[a] The first letter designations for the samples include C = Corcoran, SW = Southwest transect, X = Intensive Sampling.
The number designations refer to the Section from which the sample was collected.
The final letter designates the position in the field from which the sample was collected.
upr designates unpaved road.

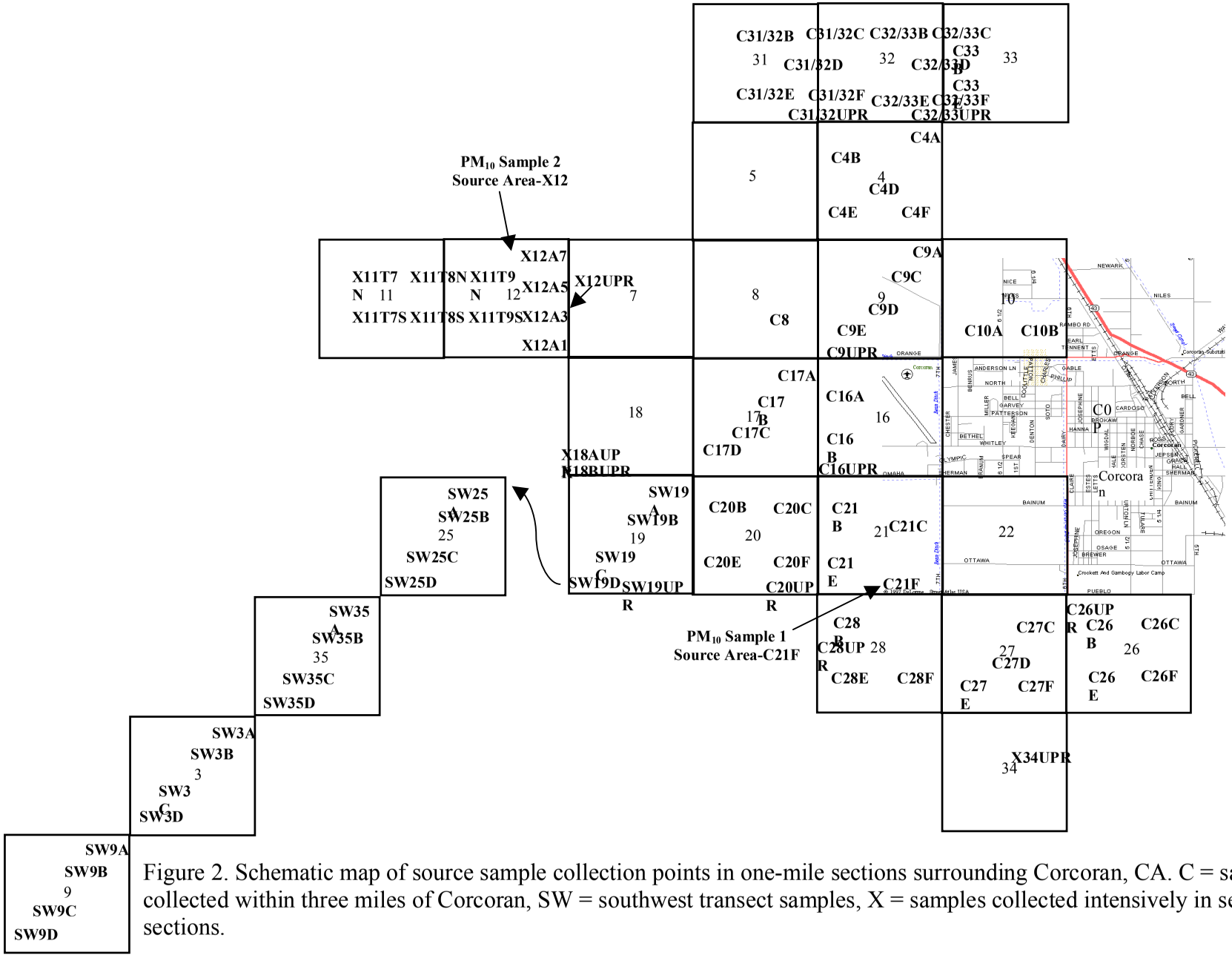


Figure 2. Schematic map of source sample collection points in one-mile sections surrounding Corcoran, CA. C = samples collected within three miles of Corcoran, SW = southwest transect samples, X = samples collected intensively in selected sections.

4.1.3 Ambient air PM_{10} at Corcoran Core Site

Twenty PM_{10} samples were collected on quartz filters from Oct 12-Nov 15, 2000 (Table 2). The lowest mass of PM_{10} (120 mg) was collected on filter Q651 and the highest mass (390 mg) was collected on filter Q646. Of the 20 filters, DNA fingerprints were obtained for 10 using routine procedures (Figure 2). The filters for which DNA fingerprints were obtained spanned two of the three peak PM_{10} events during the sampling period (Oct 19-Oct 27, 2000 and Oct 29-Nov 14, 2000). There was no correlation between the mass of PM_{10} and the production of DNA fingerprints. It is likely that the extracts of PM_{10} contained microbial DNA but at levels below the current detection limit for our analysis.

Table 2. Filter samples collected at the COP Core Site and accompanying data.

Filter Sample	Site	Start Collection	End Collection	Duration (days)	Initial Windspeed(mph)/ Direction	Final Windspeed(mph)/ Direction	Mass PM_{10} (mg)
Q644	COP	10/12/2000	10/13/2000	1	0-5, NE	0-5, W	140
Q645	COP	10/13/2000	10/16/2000	3	0-5, W	0-5, NE	360
Q646	COP	10/16/2000	10/18/2000	2	0-5, NE	0-5, NE	390
Q647	COP	10/18/2000	10/20/2000	2	0-5, NE	0-5, NW	220
Q648	COP	10/20/2000	10/22/2000	2	0-5, NW	unknown	290
Q649	COP	10/23/2000	10/25/2000	2	0-5, W	5-10, E	330
Q650	COP	10/25/2000	10/27/2000	2	5-10, E	5-10, E	220
Q651	COP	10/27/2000	10/29/2000	2	5-10, E	0-5, N	120
Q652	COP	10/30/2000	11/1/2000	2	0-5, N	0-5, W	180
Q653	COP	11/1/2000	11/3/2000	2	0-5, W	0-5, S	280
Q654	COP	11/3/2000	11/5/2000	2	0-5, S	0-5, E	280
Q655	COP	11/6/2000	11/8/2000	2	0-5, E	0-5, W	220
Q656	COP	11/9/2000	11/10/2000	1	0-5, NE	0-5, NE	150
Q658	COP	11/10/2000	11/12/2000	2	0-5, NE	unknown	170
Q659	COP	11/13/2000	11/15/2000	2	0-5, W	0-5, W	210
Q661	Field blank number 1 - remaining from static hi-vol COP site run						
Q662	Field blank number 2 - remaining from static hi-vol COP site run						
Q642	Lab blank number 2 - setup for static hi-vol COP site						
Q643	Lab blank number 1 - setup for static hi-vol COP site						

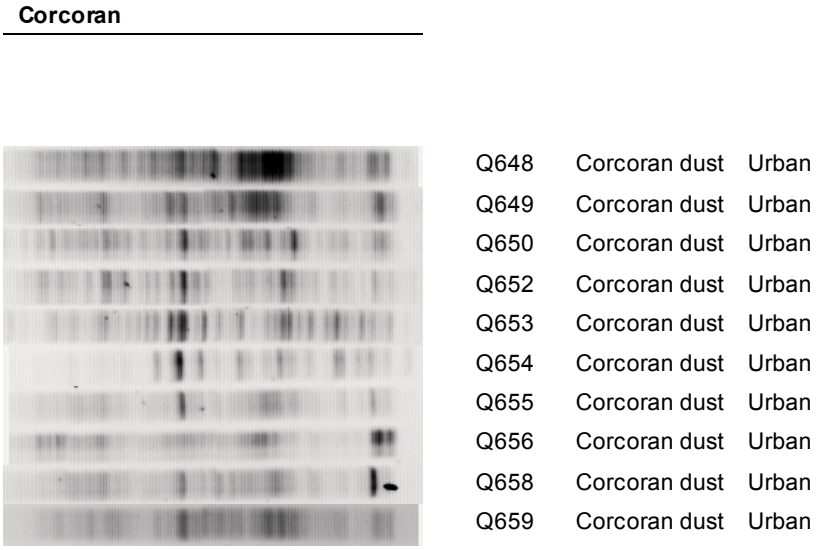


Figure 3. Microbial DNA fingerprints of ambient air PM_{10} samples

4.2 DNA Fingerprinting

4.2.1 Source sample characterization

DNA fingerprints were generated for source samples that consisted mostly of soils under cotton, wheat, hay, oats and alfalfa at different time points with respect to agricultural management. Samples were first compared within each crop type. Many of the microbial DNA fingerprints clustered by 1-mile section (as designated in Figure 2) for cotton (Figure 3), wheat (Figure 4), and oat (Figure 5). There were, however, some exceptions to these patterns.

Within cotton soils, for example, all samples within section 31 were more than 70%, and within section 20, more than 60% similar to each other. Samples in section 27 were similar to one another (>60%) with the exception of sample C27D1. Samples from section 4, on the other hand, were less than 40% similar to one another, and some of these differences may have been associated with different management practices (e.g. picking) that had recently been imposed on the soils. Two samples from section 4 (C4B2 and C4B3) were very similar to samples from section 27 despite the fact that the 2 sections were located north and south, respectively, of the town of Corcoran.

This clustering pattern indicated that soil microbial communities for the same crop within a close proximity were more similar to one another than to communities at more distant locations.

In contrast, DNA fingerprints of unpaved road samples were distinct from the fingerprints of soil microbial communities in fields adjacent to the roads. Instead, the

road samples grouped by source category rather than by location (Figure 6). This result indicated that unpaved roads have microbial communities that reflect the environmental conditions of the roads themselves and are not dominated by agricultural soil that is deposited upon the roads. Thus, it may be possible to differentiate dust emitted from roadways and dust emitted from agricultural soils.

Figure 7 shows the results of the DNA fingerprinting of soil samples collected along the southwest transect originating in Corcoran. As depicted in Figure 2, the samples are arranged in order from SW19 (nearest Corcoran), SW 25, SW 35, SW3, SW9 along the transect. With the exception of two of the alfalfa samples collected in SW 25, all samples, regardless of crop type, were greater than 65% similar to one another.

Figure 8 shows fingerprint data from soils collected in two fields (X11 and X12) where samples were collected on a more closely spaced grid than were all other samples. These 2 fields are located within sections 11 and 12, adjacent to one another (e.g., no more than 2 miles apart). There was a strong separation of the unpaved roads from all other samples, similar to the findings shown in Figure 6. There was a very large (<20% similarity) separation of the X11 (hay) from X12 (cotton) samples. Within each field, however, soil samples were ≥70% similar to one another, reflecting considerable homogeneity in microbial communities within a field. It was not possible to determine how much the similarity is due to effects of crop type or other properties associated with the locations.

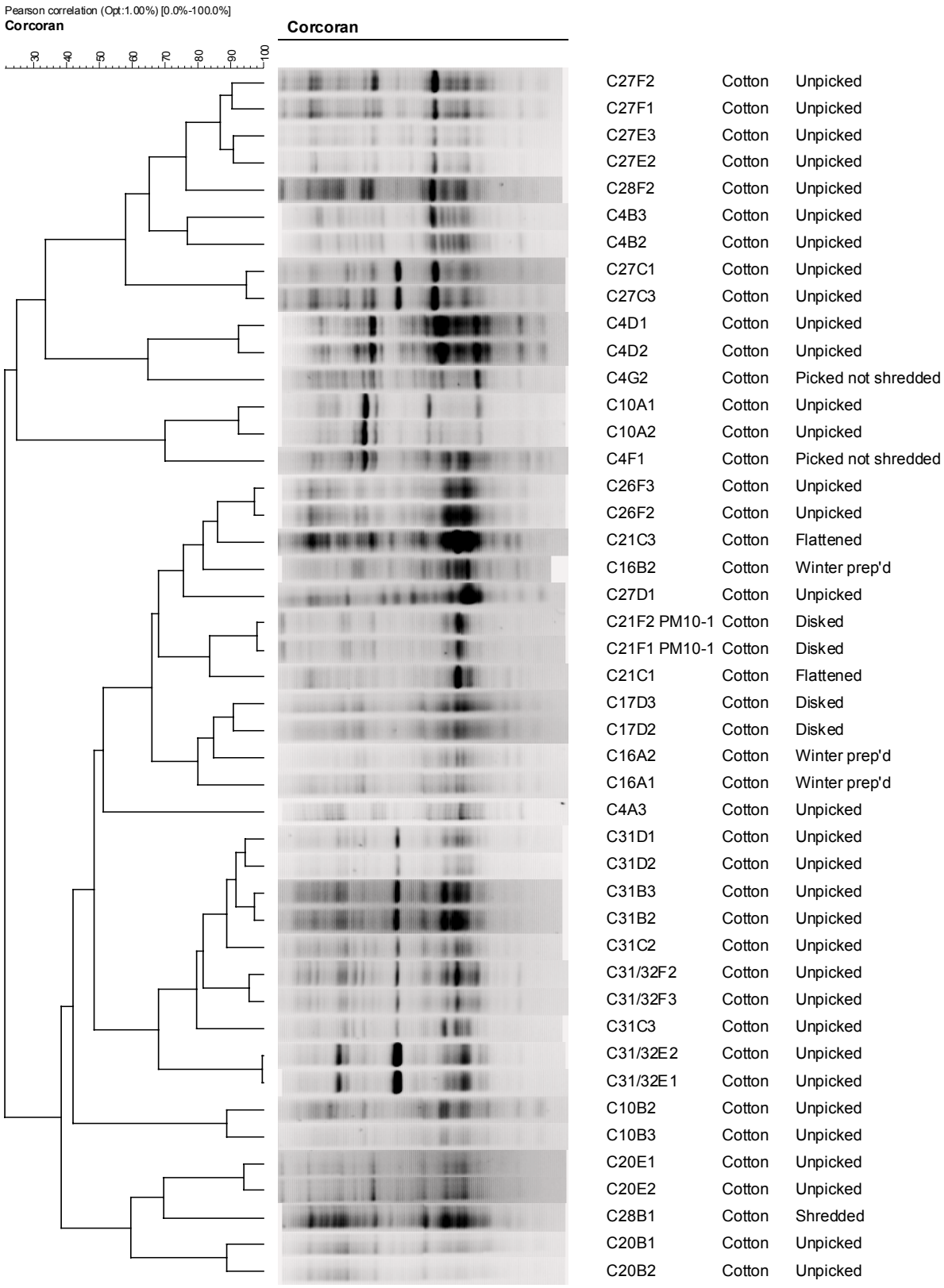


Figure 4. Cluster diagram of DNA fingerprints of cotton source samples surrounding Corcoran, CA. Profiles were analyzed by the Pearson product moment correlation and the un-weighted pair group method using arithmetic averages (UPGMA) algorithm.

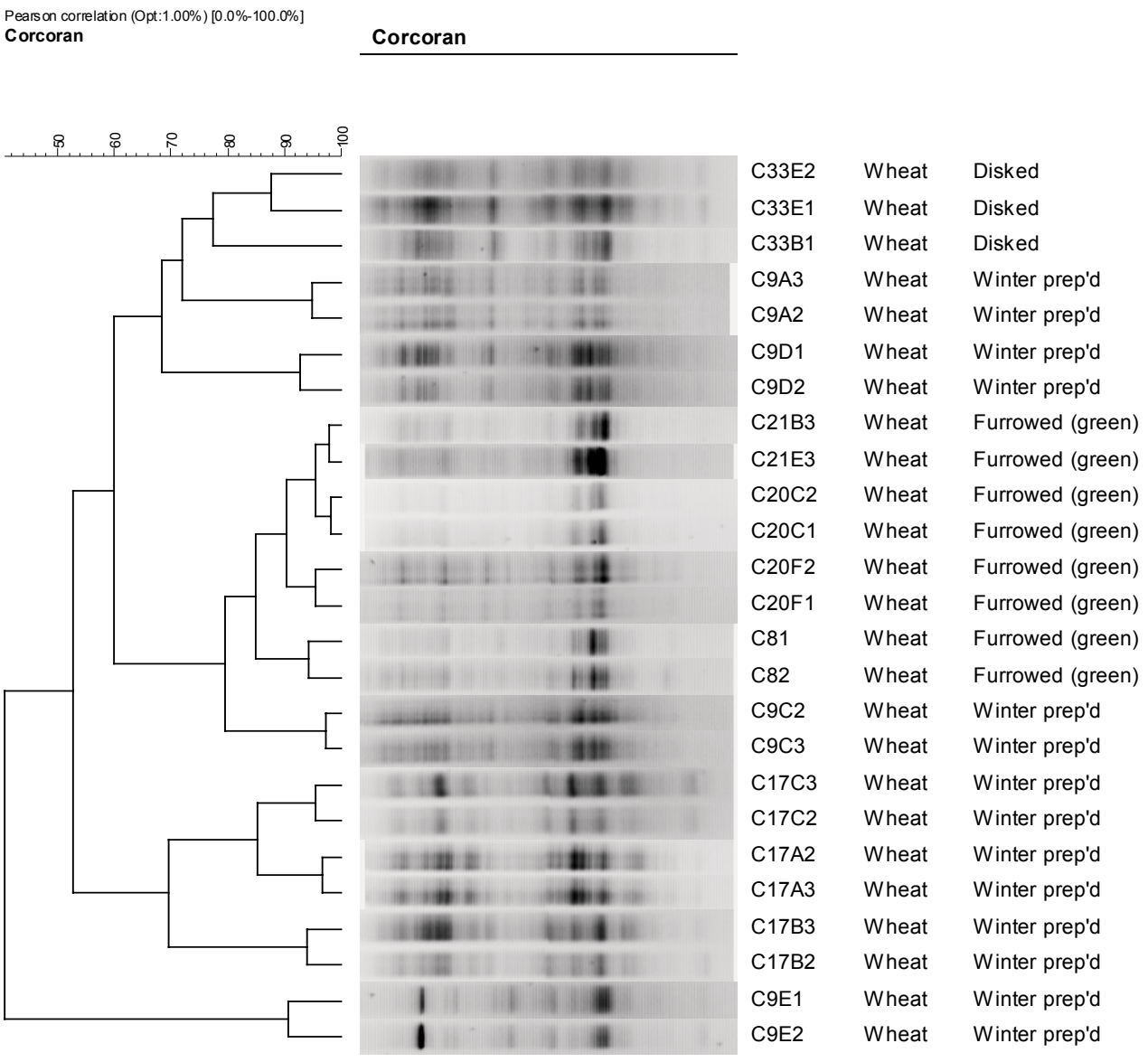


Figure 5. Cluster diagram of DNA fingerprints of wheat source samples surrounding Corcoran, CA. Profiles were analyzed by the Pearson product moment correlation and the un-weighted pair group method using arithmetic averages (UPGMA) algorithm.

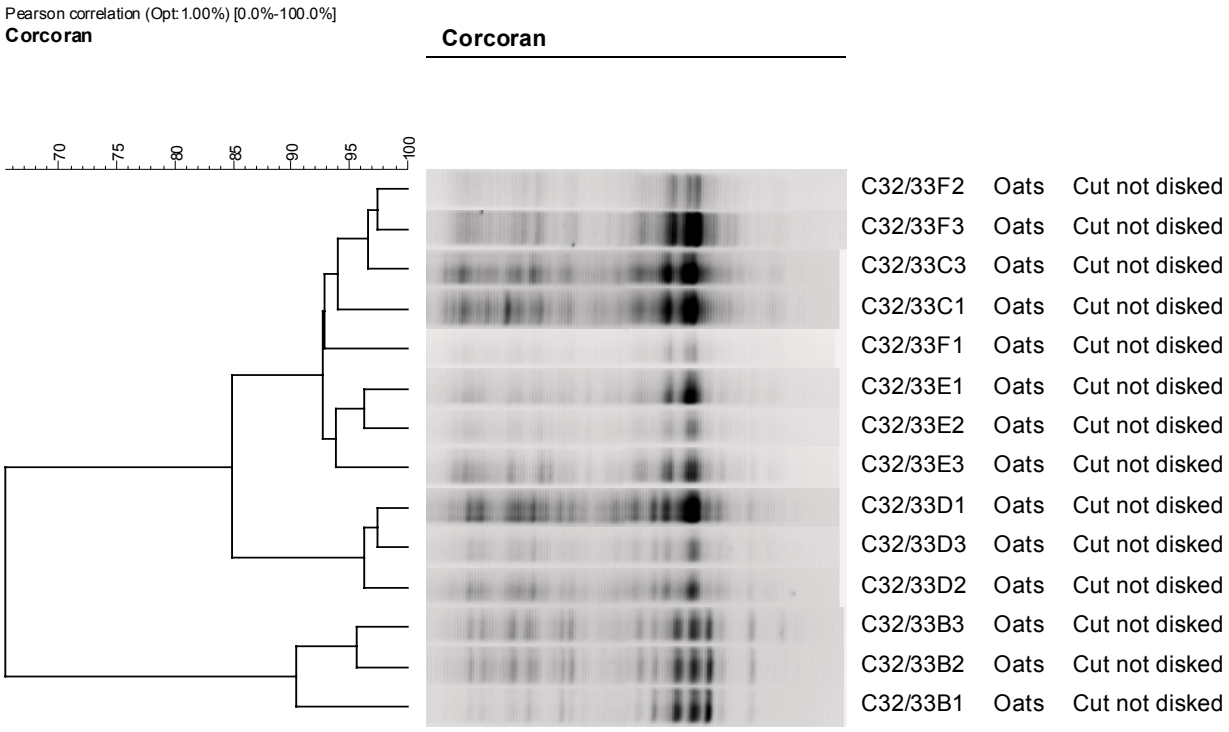


Figure 6. Cluster diagram of DNA fingerprints of oat source samples surrounding Corcoran, CA. Profiles were analyzed by the Pearson product moment correlation and the un-weighted pair group method using arithmetic averages (UPGMA) algorithm.

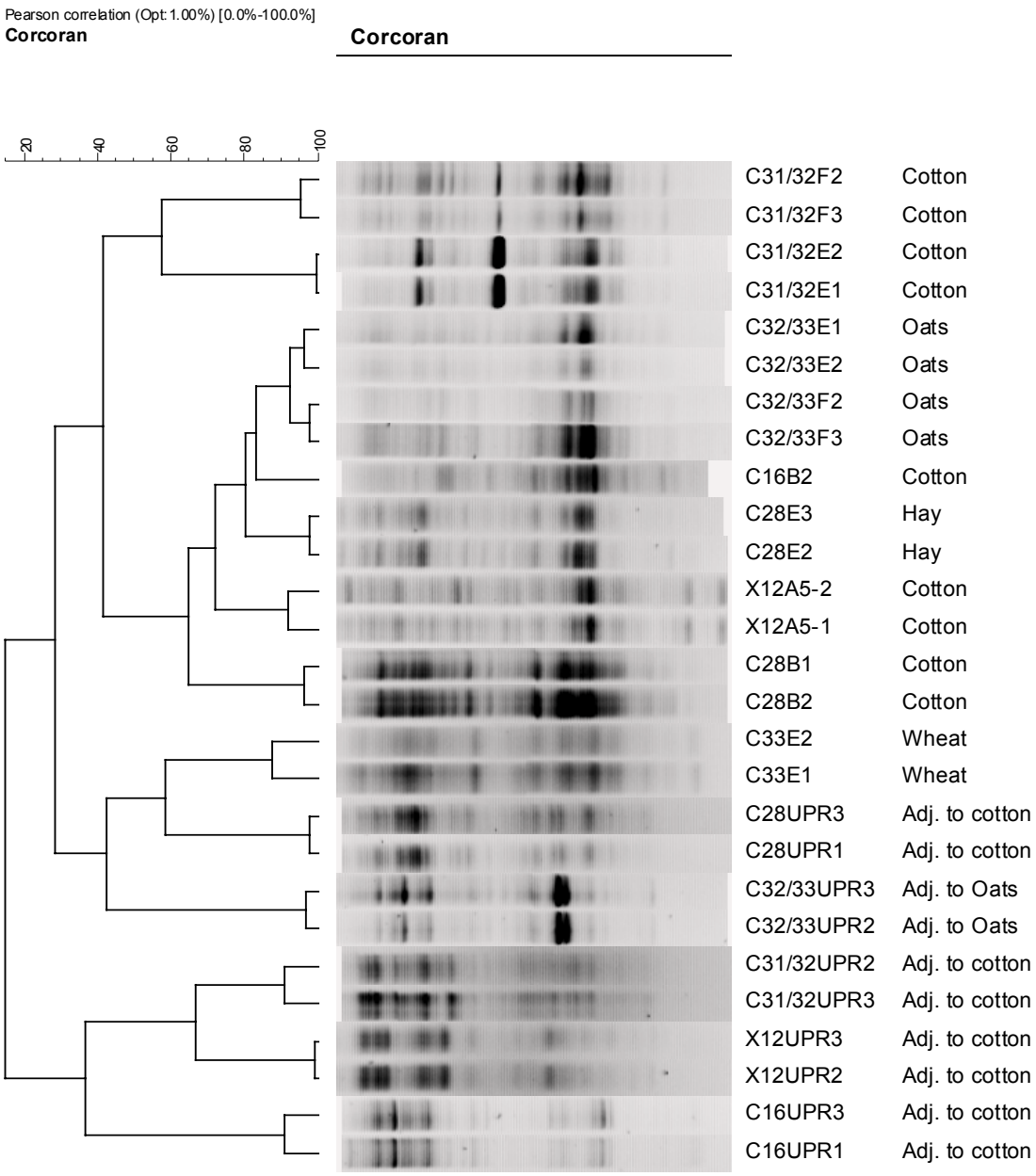


Figure 7. Cluster diagram of DNA fingerprints of unpaved road samples and the fields adjacent to the unpaved roads. Profiles were analyzed by the Pearson product moment correlation and the un-weighted pair group method using arithmetic averages (UPGMA) algorithm

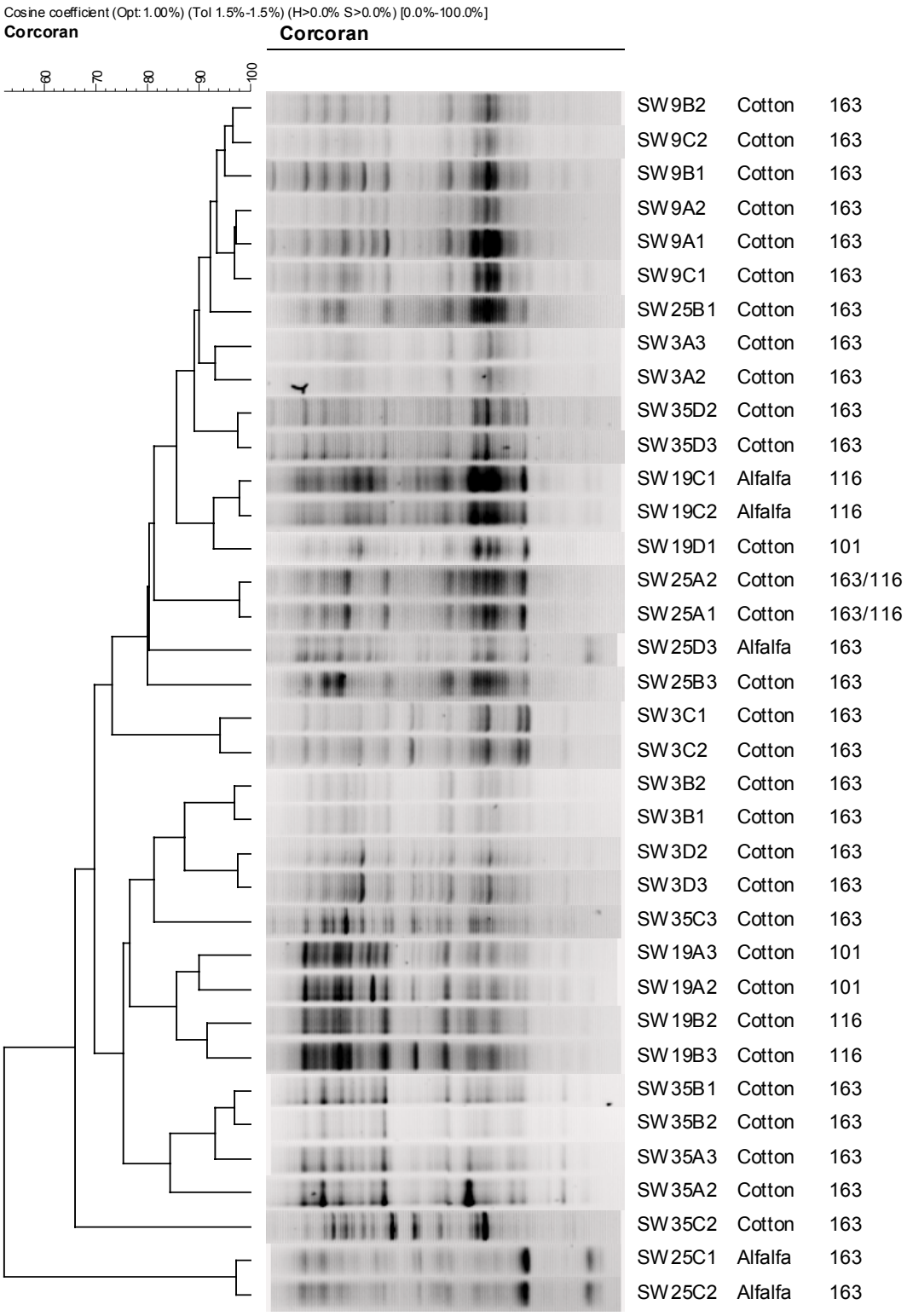


Figure 8. Cluster diagram of DNA fingerprints of southwest transect sources. Profiles were analyzed by the Pearson product moment correlation and the un-weighted pair group method using arithmetic averages (UPGMA) algorithm.

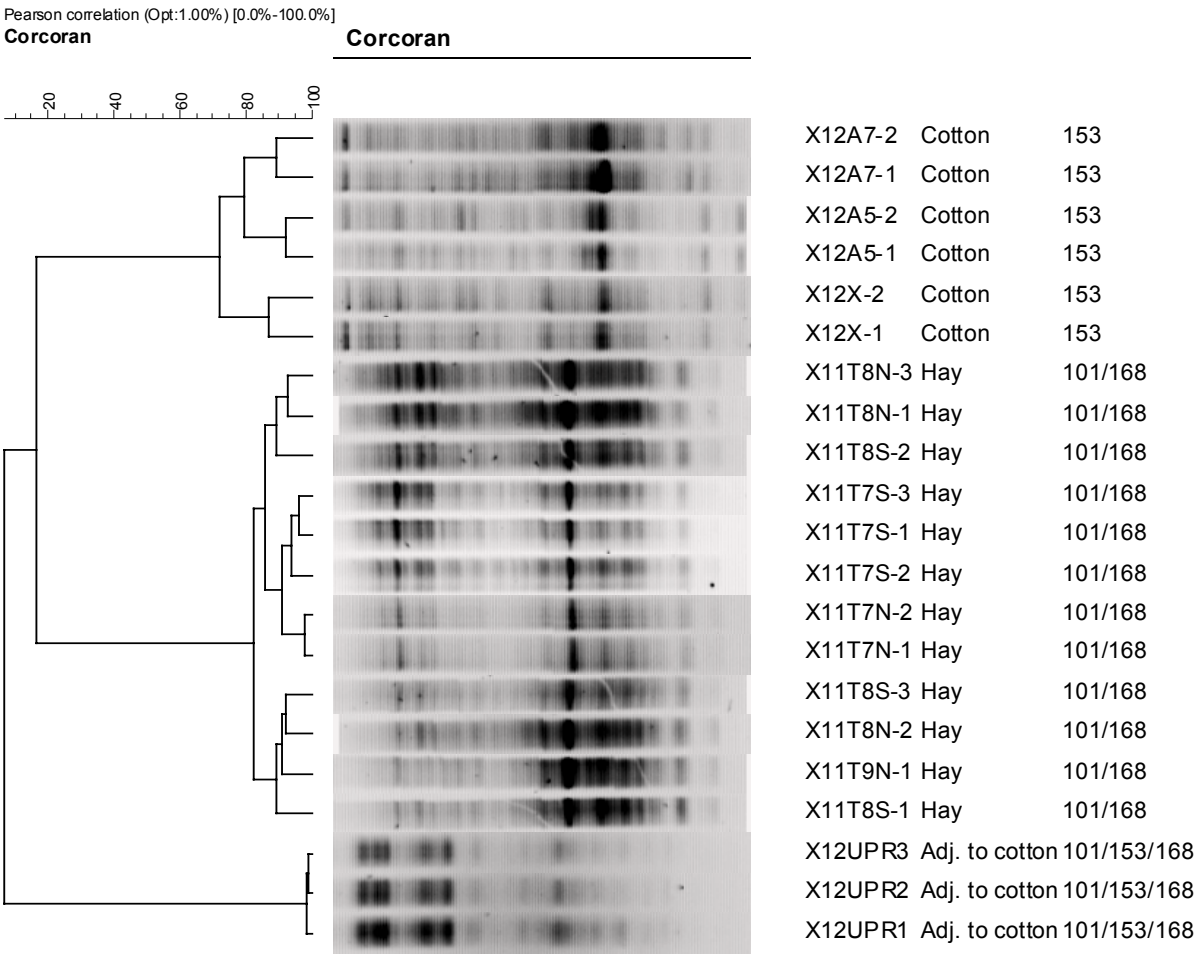


Figure 9. Cluster diagram of DNA fingerprints of intensively sampled fields. Profiles were analyzed by the Pearson product moment correlation and the un-weighted pair group method using arithmetic averages (UPGMA) algorithm.

4.2.2 Agricultural field PM_{10}

Samples of dust suspended during agricultural management practices, along with source soil samples, were collected and compared. Figure 9 shows a digital version of the DNA fingerprinting gels for the dust and source soil samples for Q601 and

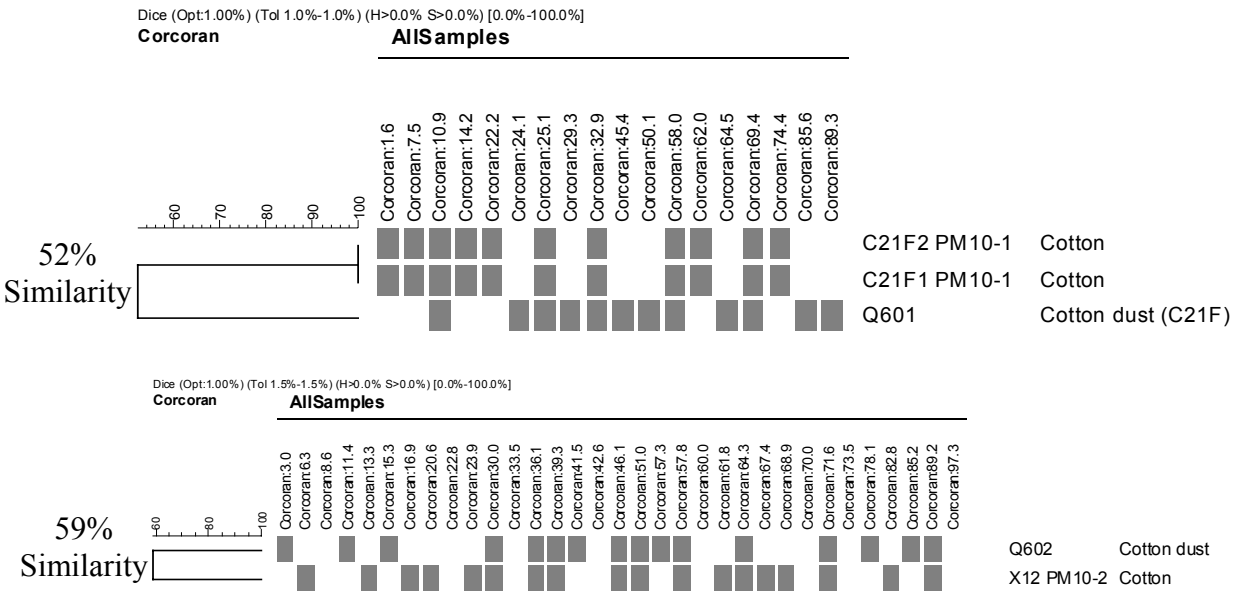


Figure 10. Cluster diagrams comparing DNA fingerprints of PM_{10} samples collected during agricultural operations to their emission source. Profiles were analyzed by the Dice coefficient and the un-weighted pair group method using arithmetic averages (UPGMA) algorithm.

Q602. The DNA fingerprint of PM_{10} sample Q601 contained five bands occurring at the same retention times (e.g., locations within the gel) as bands in the source samples (two replicates). In the case of sample Q602, there were nine bands in common between the dust and source soil samples, strongly suggesting that microorganisms in source samples were attached to suspended dust particles, captured on the quartz filters, and thus could be detected in the ambient air samples. Positive confirmation that the organisms in source soils and dust are the same requires excising the bands, cloning and sequencing them, and then comparing the DNA sequences of soil and dust bands. This effort was beyond the scope of this project.

4.2.3 Ambient Air PM_{10} at Corcoran Core Site

Nine ambient air samples of PM_{10s} dust were collected at the Corcoran Core site during the intensive sampling campaign in October, 2000. The DNA fingerprints of these

samples are first analyzed and discussed on an individual basis, or combined in small groups of related samples, and then all samples together are considered. The fingerprints of dust samples are compared to those of potential source soil samples.

PM_{10} sample Q648: This sample was collected during the two-day period immediately following the first peak PM_{10} event (10/20/2000-10/22/2000). The mass collected was 290 mg. At the onset of sampling, the wind speed was 0-5 mph from the northwest. The wind direction at the completion of sampling was not available. Sources sampled northwest from the COP anchor site include C31 (cotton), C32 (oats), C33 (wheat), C4 (cotton), C9 (wheat), and C10 (cotton). Assuming the wind continued from the northwest through the majority of the sampling period, and that agricultural operations were taking place during the sampling period in these areas, the most likely sources were determined based on comparing the banding patterns of their DNA fingerprints. According to the cluster analysis, the DNA fingerprint of Q648 is most similar to those of oat and wheat samples (Figure 10**)**. Sixteen markers were detected in PM_{10} sample Q648 by DNA fingerprinting, and four of those markers (B1-B4) were consistently detected in oat and wheat sources located northwest of the COP anchor site.

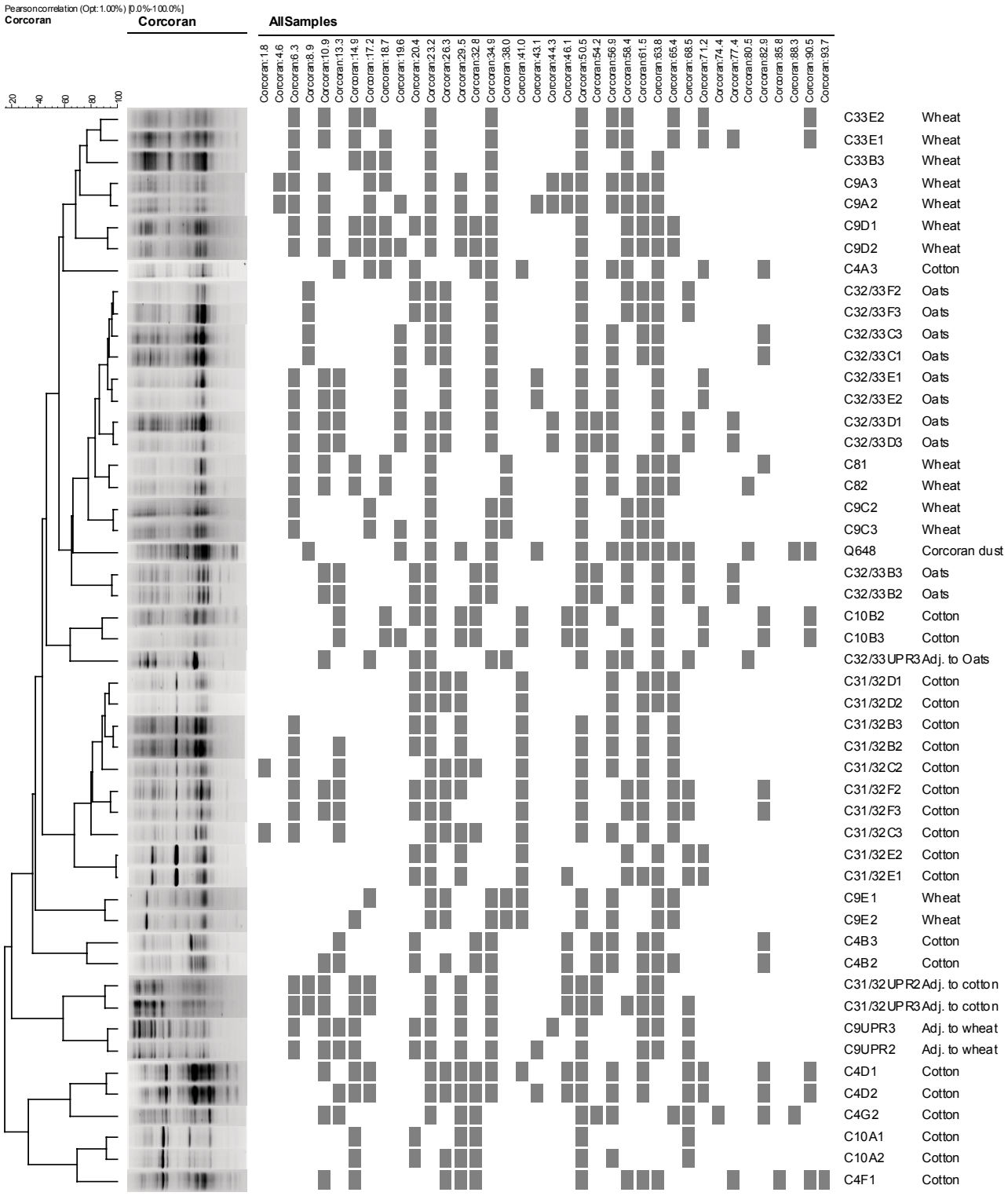


Figure 11. Cluster diagram of DNA fingerprints of PM_{10} sample Q648 (outlined with box) and sources collected northwest of Corcoran, CA. Profiles were analyzed by the Pearson product moment correlation and the un-weighted pair group method using arithmetic averages.

PM_{10} samples Q649, Q650, and Q651: These samples were collected over a period of five days at the peak of, and immediately following the second PM_{10} event. This period of time concluded with several days of rainfall (10/23/2000-10/29/2000). The masses of dust collected were 330 mg (Q649), 220 mg (Q650), and 120 mg (Q651). At the onset of sampling, the wind speed was 0-5 mph from the west, but shifted to an easterly direction with an increased speed to 5-10 mph for at least 75% of the sampling period. Source sampling was not conducted east of the COP anchor site, as prevailing winds from that direction were not typical for the study period. Thus, it was not possible to relate the DNA fingerprints of potential sources to these dust samples.

Several observations can be made regarding the DNA fingerprints of this set of dust samples. A total of ten DNA bands were detected in the DNA fingerprint of filter sample Q649, whereas 18 were detected in sample Q650 (Figure 11**)**. Of the total number of bands detected in the two samples, two were unique to Q649 and ten were unique to Q650. The two samples shared a 46% similarity. The presence of additional markers in sample Q650, despite the lower filter mass, may indicate the contribution of dust from immediately adjacent sources (e.g., within Corcoran) because the rainfall event dampened the production and transport of dust from agricultural sources valley wide. Potential local sources of dust include the GRA and YOD Corcoran focus sites located east of the COP anchor site. An atypically high mass of dust (over 30 $\mu g/m^3$) was recorded for GRA on 10/29/2000 and may have been derived from these local sources. It is possible that some of the DNA bands detected in Q650 were markers for local sources of dust; this hypothesis, however, requires further testing.

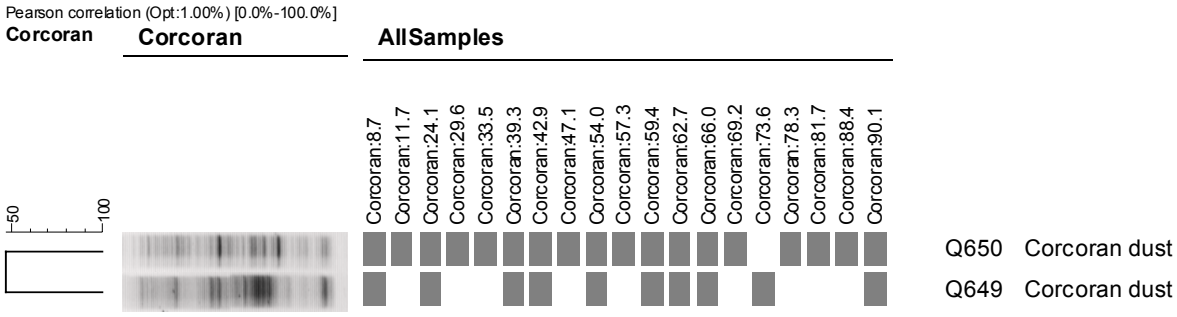


Figure 1. DNA fingerprints of PM_{10} samples Q649 and Q650.

PM_{10} samples Q652, Q653, and Q654: These samples were collected over a six day period encompassing a period of time before, at the peak of, and immediately following the third PM_{10} event. This third event concluded with several days of rainfall (10/30/2000-11/5/2000). Samples Q652, Q653, and Q654 were collected two days apart from one another, on 11/1, 11/3, and 11/5, respectively. The masses of dust collected were 180 mg (Q652), 220 mg (Q653), and 220 mg (Q654). During the collection period, the wind shifted at least four times and originated from the north, west, south, or east. Thus, it was not possible to isolate potential sources for comparison to dust samples.

Based on a whole-pattern analysis of the DNA fingerprints, sample Q652 was distinct from Q653 and Q654 (the latter two samples showing a 73% similarity**)** (Figure 12). The lowest number of bands was detected in sample Q654, which was collected at the peak of the third PM_{10} event. A similar observation of fewer numbers of bands was also made in samples collected during the second PM_{10} event (filter Q649 from 10/23/00-10/25/00). It is possible that samples collected during peak PM_{10} events are dominated by one or a very few sources and thus a lower number of bands are detected than when there are contributions from multiple sources.

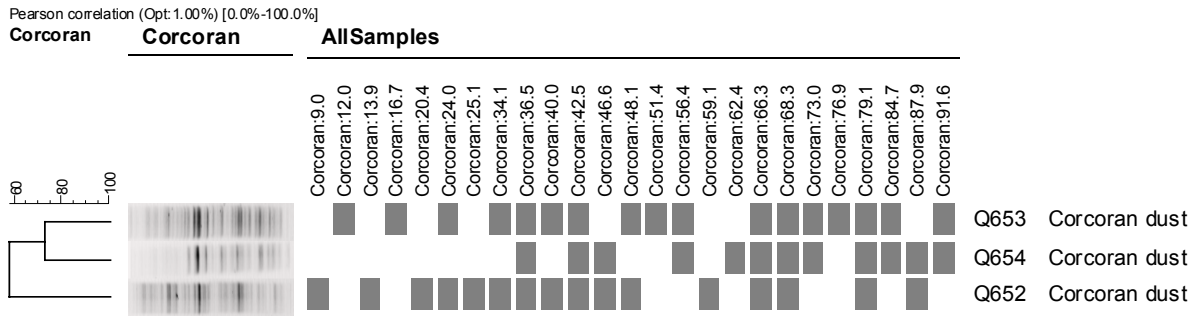


Figure 2. DNA fingerprints of PM_{10} samples Q652, Q653, and Q654.

PM_{10} sample Q655: This dust sample was collected during a two-day period in which PM_{10} concentrations at Corcoran exceeded 50 $\mu g/m^3$, but were variable valley wide (11/6/2000-11/8-2000). At the onset of sampling, the wind speed was 0-5 mph from the east. At the completion of sampling the wind speed was 0-5 mph from the west. The mass of dust collected was 220 mg. Potential sources sampled included samples west of C16 (cotton) and C17 (wheat) and unpaved roads adjacent to cotton or alfalfa, also west of the COP anchor site. Based on DNA fingerprint analysis, the Q655 sample was quite distinct (< 20% similar) from the possible source samples (Figure 13). The Q655 sample contained two bands not present in any of the source samples; however the remaining 11 bands appeared to be present in some of the potential source soils. A possible reason for the lack of correlation between the dust and any one of the source samples could be that multiple sources contributed some, but not all of their bands, to the PM_{10} DNA fingerprint. Thus Q655 represents a composite of multiple sources.

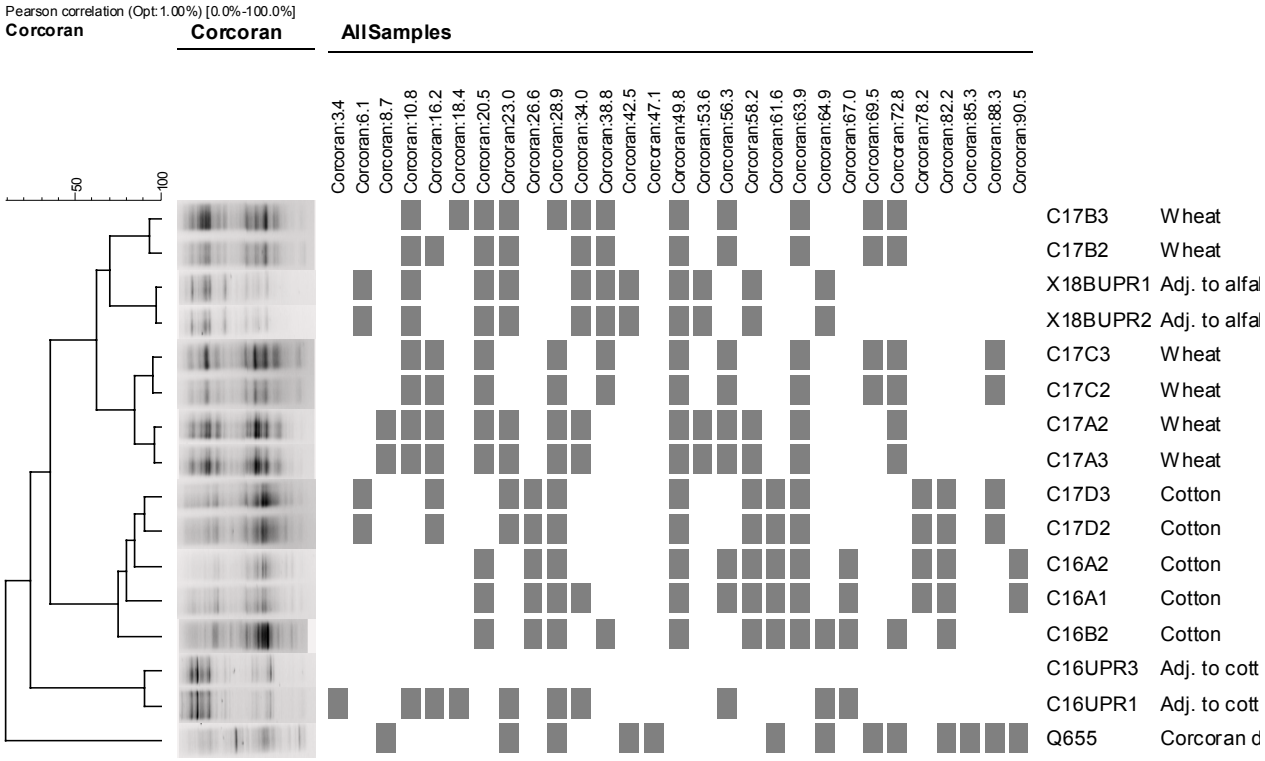


Figure 3. Cluster diagram of DNA fingerprints of PM_{10} sample Q655 and sources collected west of Corcoran, CA. Profiles were analyzed by the Pearson product moment correlation and the un-weighted pair group method using arithmetic averages (UPGMA) algorithm.

PM_{10} sample Q658 and Q659: These dust samples were collected during the latter part of the sampling campaign, on 11/12/2000 and 11/15/2000, both after 2 days of sampling. The masses of dust collected were 170 and 210 mg, respectively. The wind direction was primarily NE for Q658 and W for Q659. The samples showed a 60% similarity, with the Q658 having far fewer DNA bands than Q659.

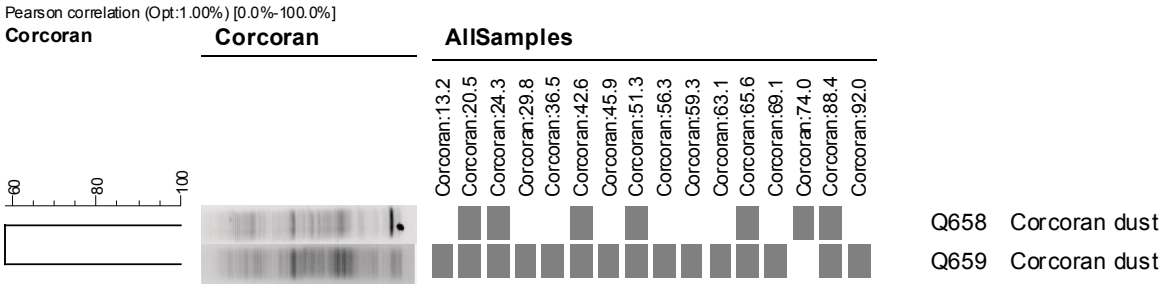


Figure 4. DNA fingerprints of PM_{10} samples Q658 and Q659.

Comparison of the entire set of PM_{10} samples. The DNA fingerprints of all PM_{10} samples were compared to evaluate common trends and changes with time. The cluster analysis grouped the DNA fingerprints of PM_{10} samples into three main associations (Figure 15). The first group (Q655, Q659, Q650, and Q648) was associated with periods of time during which PM_{10} levels were declining, or were at their lowest concentrations. The samples Q656, and Q658, comprising the second group, were associated with the between peak PM_{10} levels on 11/9/2000 and 11/13/2000. The third and most distinct group (Q652-Q654) was associated with high concentrations of PM_{10} that peaked on 11/4/2000. From this analysis, we can conclude that PM_{10} fingerprints vary with changes in the contributions of different sources.

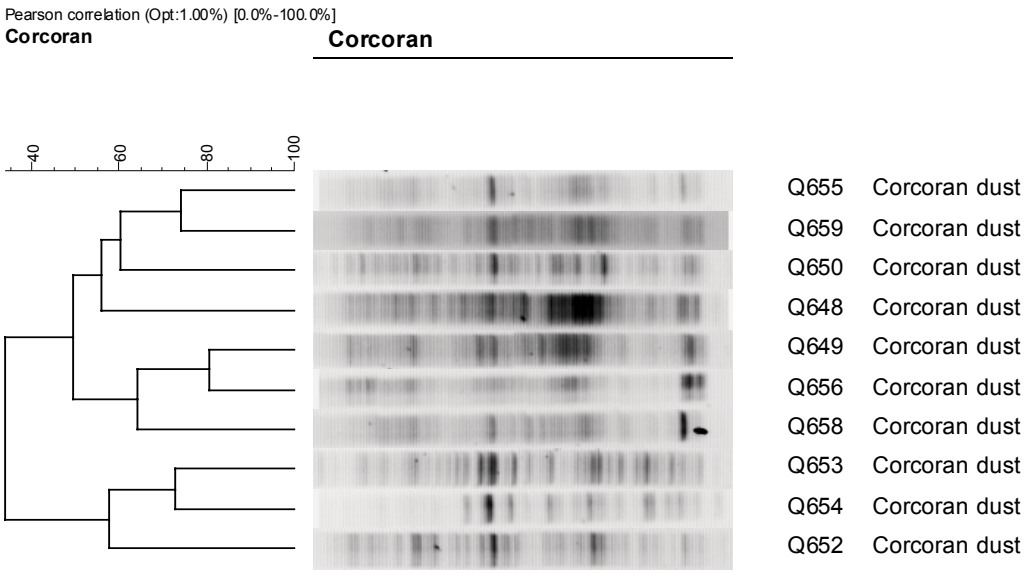


Figure 5. Cluster diagram of all DNA fingerprints of PM_{10} samples collected at the COP Core Site in Corcoran, CA. Profiles were analyzed by the Pearson product moment correlation and the un-weighted pair group method using arithmetic averages (UPGMA) algorithm.

5.0 Major Conclusions

The Corcoran study is the first field study to evaluate the efficacy of using DNA fingerprinting to help identify sources of PM_{10} collected in Hi-Vol samplers. Although aspects of the study were limited by rainfall during key sampling periods, which resulted in reduced dust accumulation on filters, a number of conclusions can be drawn from this work. These are as follows:

1. We were successful in extracting and analyzing DNA from PM_{10} collected on Hi-Vol sampler filters at the Corcoran stationary site and from filters collected in the field during agricultural operations. The filters at the stationary site, which collected PM_{10} over 2 day sampling periods, contained masses of dust from 120 to 390 mg. These amounts are sufficient to obtain enough DNA to permit PCR-based fingerprint analysis.

- *Microbial DNA was routinely detected in PM_{10} on sterilized quartz filters. Although the absolute lower limits of detection were not determined in this study, the weakest signal (e.g fewest number of bands) was observed from the filter with the lowest sample mass (120 mg). Thus this amount may be approaching the lower limit of detection.*
- *Low masses of DNA coincided with days during, and following, rainfall events.*
- *Although the mass of DNA extracted from filters is lower than that extracted from source samples, the DNA fingerprints of PM_{10} and sources were similar with respect to the amount of information (e.g., number of bands per fingerprints) they contained. This made it possible to compare dust and potential source samples.*

2. The DNA fingerprints of dust on filters changed over time and varied with different peak PM_{10} events. Thus, there was no unique dust fingerprint common to all dust samples. Dust samples were more similar to their potential sources than to one another. This provided support for the idea that DNA fingerprinting is potentially useful for relating PM_{10} to potential sources.

3. In some, but not all cases, there were similarities between dust samples collected at the Corcoran stationary site and from some of the sources. For some of the samples, a subset, but not all, bands within their fingerprints appeared to be the same. In some cases, source samples upwind of the Corcoran stationary site were not collected because the wind direction did not follow predicted patterns. In all cases, there was no opportunity for replication in sample collection and thus hindered the ability to detect small differences and to draw statistically-based conclusions.

4. We collected a library of DNA fingerprints for potential sources samples from agricultural soils and unpaved roads surrounding Corcoran. We could differentiate source samples from one another, but there did not appear to be a strong relationship between source fingerprint and land use or crop type (in contrast to findings in our previous ARB contract).

6.0 RECOMMENDATIONS

The weather conditions during the intensive sampling period in Corcoran prevented the collection of a large number of dust samples. While the source sampling was planned to collect samples from direction of prevailing winds in typical years, this year was not typical. Thus it is difficult to identify potential sources within a 2-3 mile radius of Corcoran the 2-day filter collection runs. If such a study was to be repeated, a sampling plans with a broader network of support staff will be required to adequately assess the potential origins of dust collected at the anchor site in Corcoran and develop links between sources and filters. It would help such an effort to include aerial or other observations to ascertain the areas of primary emission during a particular filter sampling period.

Another recommendation for further studies would be to explicitly test the hypothesis that fugitive dust sources can be characterized by location. To do this, a study designed to examine the extent of variability within what is designated as a "source" is required. It is possible that the arbitrary designation of "field" may not be the appropriate unit of delineation. It is possible that soil mapping on the scale of conventional soil surveys or of the scale in the STATSGO database would be more appropriate. Connections between particle types and DNA fingerprints have already been observed with the map unit identification system in the STATSGO database. We recommend this as a possible framework for organizing future sample collection.

With regard to the DNA-based methods, we recommend that further studies focus on potential specific biomarkers (e.g., single DNA bands) that are indicators of a source. These biomarkers could be quantified using other emerging molecular tools and would greatly simplify the data analysis currently required in the evaluation of entire DNA fingerprints.

7.0 REFERENCES

Borneman, J. and E. W. Triplett (1997). Molecular microbial diversity in soils from eastern Amazonia: Evidence for unusual microorganisms and microbial population shifts associated with deforestation. Applied and Environmental Microbiology **63**: 2647-2653.

Dice, L. R. (1945). Measures of the amount of ecologic association between species. Ecology **23**: 297-302.

Haene, B. G., K. Jaeger and H. G. Drexler (1993). The Pearson product-moment correlation coefficient is better suited for identification of DNA fingerprint profiles than band matching algorithms. Electrophoresis **14**: 967-972.

Jongman, R. H., C. J. F. ter Braak and O. F. R. Van Tongeren (1995). Data Analysis in Community and Landscape Ecology. Cambridge, Cambridge University Press.

Palumbi, S. R. (1996). Nucleic acids II: The polymerase chain reaction. Molecular Systematics. D. M. Hillis, C. Moritz and B. K. Mable. Sunderland, MA, Sinauer Associates**:** 205-247.

Pearson, K. (1926). On the coefficient of racial likeness. Biometrika **18**: 105-117.

Rademaker, J. L. W. and F. J. De Bruijn (1997). Characterization and classification of microbes by rep-PCR genomic fingerprinting and computer assisted pattern analysis. DNA markers: Protocols, Applications, and Overviews. G. Caetano-Anolles and G. P. M. New York, Wiley-Liss, Inc.**:** 151-171.

Sneath, P. H. A. and R. R. Sokal (1973). Numerical Taxonomy; The Principles and Practice of Numerical Classification. San Francisco, W. H. Freeman.

ter Braak, C. J. F. (1994). Canonical community ordination. Part I: Basic theory and linear methods. Ecoscience **1**: 127-140.

The Institute for Genomic Research (TIGR) (2001). Comprehensive Microbial Resource, Microbial Genome Program, U. S. Department of Energy.

APPENDIX A DNA Extraction and Analysis

DNA from soil microorganisms was extracted and purified from 500 mg soil, from filter blanks, and from filters containing the PM_{10} dust samples with the FastDNA™ Spin Kit for soil and the FastPrep Instrument (Bio 101, Inc., Vista, CA) according to the manufacturer's guidelines. An empty extraction tube was included in each extraction as a blank. Four filter sections were extracted individually, combined, and concentrated. The PCR was conducted on each of the resulting filter extracts, on extracts of filter blanks, and on an extraction blank.

The polymerase chain reaction (PCR) targeting the Intergenic Transcribed Spacer (ITS) region was used to obtain DNA fingerprints of soil microbial communities. As mentioned previously, PCR primers were designed to complement either bacterial DNA. PCR amplification was directed by primers 1406f, 5'-TGYACACACCGCCCGT-3' (Universal, 16S rRNA gene) and 155r, 5'-GGGTTBCCCCATTCRG-3' (bacterial-specific, 23S rRNA gene) (Borneman and Triplett, 1997). Replicate PCRs were completed for both primer sets for all samples with serial dilutions of DNA extracts that yielded 1-4 ng of DNA template. PCR optimization was completed following the guidelines offered by Palumbi (1996) using DNA extracted from *Bacillus subtilis* (ATCC# 6051)*, Escherichia coli* (ATCC# 10798), and *Saccharomyces cervisiae* (ATCC# 204680). DNA extracts of these species were also used as positive and negative control DNA for PCR. As their compete genomic sequences are available (The Institute for Genomic Research (TIGR), 2001), the numbers and sizes of their ITS regions are known. The resolution of appropriately sized bands from these control strains confirmed optimal PCR (and electrophoresis) conditions both initially and throughout the study.

The 50 µl reaction mixture, consisting of 25 pmol of each primer, 200 µM each dNTPs, 1X PCR buffer, 2.0 mM $MgCl_2$, and 1.5 U AmpliTaq Gold™ DNA polymerase (Perkin Elmer Applied Biosystems, Foster City, CA) was combined with 2 µl of the 2-fold dilutions of template DNA. In addition, positive and negative control DNA (see above), solutions from DNA extraction blanks, and sterilized nanopure water as a PCR blank were included in each PCR. After a pre-incubation step to activate the AmpliTaq Gold (95° C for 10 min), thermocycling consisted of 30 cycles of denaturation at 94° C (30 s), annealing at either 60° C (30 s), extension at 72° C (1 min), and a final extension at 72° C (10 min). All PCR products were examined by agarose gel electrophoresis with ethidium bromide stain to estimate the volumes of PCR products to load in polyacrylamide gels.

PCR product solutions (4-10 µl) were loaded in 4% polyacrylamide/1X TBE gels and electrophoresed in the DCode™ System (Bio-Rad Laboratories, Inc., Hercules, CA) at 150V (6.8 V cm^{-1}) for 3.5 hours at 25° C. After staining with 0.01% SYBR Green (BioWhittaker Molecular Applications, Rockland ME) for 30 min, gels were illuminated with UV light for image capture with a charge-coupled-device (CCD) camera equipped with a 520 nm bandpass filter (Corion Corp., Franklin, MA). Two DNA fingerprints, representing two dilutions of template DNA for a single sample, were included in the image and statistical analyses.

APPENDIX B Image Processing Points for DNA Fingerprints

Overview

To compare DNA fingerprints of more than 16 samples (the number of lanes within a single gel), a standardized system of producing and processing multiple gel images is required. This task was completed with the acquisition and use of a state-of-the-art software package for processing DNA fingerprints called Gelcompar II (Applied Maths, Kortrijk, Belgium). The specific settings used in image analysis of DNA fingerprint patterns are detailed below.

DNA fingerprint gel image processing

The settings used to process inverted gel images included 35 point, 3 node image strip extraction for the raw data, an averaging thickness of 13 points with 3 nodes for curve extraction, a rolling ball background subtraction of 10% (based on spectral curve analysis of all gels with the Fourier method), and arithmetic average and least square filtering with a cutoff below 1% and a power of 2.0. The vertical dimension of the gel consisted of 312 pixels. All gels lanes were normalized to a common reference pattern, a 20 bp DNA ladder (BioWhittaker Molecular Applications, Rockland ME), which was loaded in at least 4 positions on every gel. To designate bands in the profiles, the band search filter was set to a minimum profiling of 5% relative to maximum value, a minimum area of 0.5%, and a shoulder sensitivity of two. The positions of bands within the fingerprints were located by fitting the peaks of intensity by regression (cubic spline fit with logarithmic dependence) against the migration of bands in the 20 bp DNA ladder. Band positions were only derived for a specific region on the gel, namely within the bounds of the 20 bp ladder (1000-300 bp range). Quantitative values for the bands in a profile were obtained by integrating one-dimensional band areas (Gaussian fit). The array of bands selected by the software in each lane was inspected to remove false information (e.g., small illuminated spots from dust that met band choosing criteria of the software), or to include appropriate information (e.g., visually detected bands that did not meet the band choosing criteria of the software), and to confirm that one-dimensional band areas were integrated accurately.